METHOD FOR LEARNING AND COMBINING GLOBAL AND LOCAL REGULARITIES FOR INFORMATION EXTRACTION AND CLASSIFICATION

ABSTRACT OF THE DISCLOSURE

A method is provided for information extraction and classification which combines aspects of local regularities formulation with global regularities formulation. A candidate subset is identified. Then tentative labels are created so they can be associated with elements in the subset that have the global regularities, and the initial tentative labels are attached onto the identified elements of the candidate subset. The attached tentative labels are employed to formulate or "learn" initial local regularities. Further tentative labels are created so they can be associated with elements in the subset that have a combination of global and local regularities, and the further tentative labels are attached onto the identified elements of the candidate subset. Each new dataset is processed with reference to an increasingly-refined set of global regularities, and the output data with their associated confidence labels can be readily evaluated as to import and relevance.

PA 3069250 v1KRA